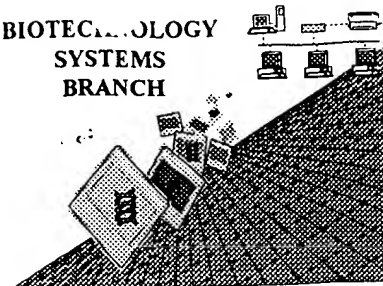


## **RAW SEQUENCE LISTING** **ERROR REPORT**

04004-30-01, 0280  
BIOTECHNOLOGY  
SYSTEMS  
BRANCH



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/821,160  
Source: OIPF  
Date Processed by STIC: 4/10/2001

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/821,160

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2        Wrapped Aminos      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3        Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4        Misaligned Amino Acid      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs  
Numbering      between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5        Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6        Variable Length      Sequence(s)        contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and  
indicate in the (ix) feature section that some may be missing.
- 7        PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid  
sequence(s)       . Normally, PatentIn would automatically generate this section from the  
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section  
to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>  
sections for Artificial or Unknown sequences.
- 8        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence:  
(OLD RULES)      (2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence.  
(NEW RULES)      <210> sequence id number  
<400> sequence id number  
000
- 10        Use of n's or Xaa's      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11        Use of "Artificial"      Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules.  
(NEW RULES)      Valid response is Artificial Sequence.
- 12        Use of <220>Feature      Sequence(s)        are missing the <220>Feature and associated headings.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13        PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted  
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,160

DATE: 04/10/2001

TIME: 15:08:23

Input Set : A:\PTO.txt

Output Set: N:\CRF3\04102001\I821160.raw

Does Not Comply  
Corrected Diskette Needed

pg 1-3

3 <110> APPLICANT: Yu, Zhongping  
 5 <120> TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING POLYPEPTIDES AND  
 NUCLEIC ACID  
 6 MOLECULES  
 8 <130> FILE REFERENCE: SEL-00104.P.1-US  
 10 <140> CURRENT APPLICATION NUMBER: US/09/821,160  
 10 <141> CURRENT FILING DATE: 2001-03-29  
 10 <150> PRIOR APPLICATION NUMBER: US 60/156,990  
 11 <151> PRIOR FILING DATE: 1999-11-01  
 13 <150> PRIOR APPLICATION NUMBER: US 60/178,420  
 14 <151> PRIOR FILING DATE: 2000-01-27  
 16 <150> PRIOR APPLICATION NUMBER: PCT/US00/26511  
 17 <151> PRIOR FILING DATE: 2000-09-27  
 19 <160> NUMBER OF SEQ ID NOS: 15  
 21 <170> SOFTWARE: PatentIn version 3.0  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 46  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Homo sapiens  
 28 <400> SEQUENCE: 1  
 29 gcgaagctta tataaggtac caggaggtga accatggcag ccggga 46  
 32 <210> SEQ ID NO: 2  
 33 <211> LENGTH: 62  
 34 <212> TYPE: DNA  
 35 <213> ORGANISM: Homo sapiens  
 37 <400> SEQUENCE: 2  
 38 gcgtctagat agtcagggc cctgaaaata caggttttcg ctcttagcag acattggaag 60  
 40 aa 62  
 43 <210> SEQ ID NO: 3  
 44 <211> LENGTH: 25  
 45 <212> TYPE: DNA  
 46 <213> ORGANISM: Artificial  
 48 <220> FEATURE:  
 49 <223> OTHER INFORMATION: Synthetic sequence including XbaI site  
 51 <400> SEQUENCE: 3  
 52 cgctctagac taggttattg gaaaa 25  
 55 <210> SEQ ID NO: 4  
 56 <211> LENGTH: 24  
 57 <212> TYPE: DNA  
 58 <213> ORGANISM: Artificial  
 60 <220> FEATURE:  
 61 <223> OTHER INFORMATION: Synthetic sequence including HindIII site  
 63 <400> SEQUENCE: 4  
 64 cgcaagctta ctgtttcctg tgtg 24  
 67 <210> SEQ ID NO: 5  
 68 <211> LENGTH: 38  
 69 <212> TYPE: DNA  
 70 <213> ORGANISM: Bacteriophage T7

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,160

DATE: 04/10/2001

TIME: 15:08:23

Input Set : A:\PTO.txt

Output Set: N:\CRF3\04102001\I821160.raw

```

72 <400> SEQUENCE: 5
73 agtgggtacct aatacgactc actataggag ctggaagg          38
76 <210> SEQ ID NO: 6
77 <211> LENGTH: 52
78 <212> TYPE: DNA
79 <213> ORGANISM: Bacteriophage T7
81 <400> SEQUENCE: 6
82 tcaccatggt ggcctcgaag tgtgcttgcc tatacgttgc cttcgagctc ct          52
85 <210> SEQ ID NO: 7
86 <211> LENGTH: 26
87 <212> TYPE: DNA
88 <213> ORGANISM: Influenza virus
90 <400> SEQUENCE: 7
91 ccagaattct acccatacga tgttcc          26
94 <210> SEQ ID NO: 8
95 <211> LENGTH: 26
96 <212> TYPE: DNA
97 <213> ORGANISM: Influenza virus
99 <400> SEQUENCE: 8
100 tgcctcgagc tagcactgag cagcgt          26
103 <210> SEQ ID NO: 9
104 <211> LENGTH: 105
105 <212> TYPE: DNA
106 <213> ORGANISM: Influenza virus
108 <400> SEQUENCE: 9
109 ttttaccat acgatgttcc tgactatgcg ggctatccct atgacgtccc ggactatgca          60
111 ggatcctatc catatgacgt tccagattac gctgctcagt gctag          105
114 <210> SEQ ID NO: 10
115 <211> LENGTH: 19
116 <212> TYPE: DNA
117 <213> ORGANISM: Artificial
119 <220> FEATURE:
120 <223> OTHER INFORMATION: Synthetic sequence designated GST-F459
122 <400> SEQUENCE: 10
123 tctatggcca tcatacggt          19
126 <210> SEQ ID NO: 11
127 <211> LENGTH: 18
128 <212> TYPE: DNA
129 <213> ORGANISM: Artificial
131 <220> FEATURE:
132 <223> OTHER INFORMATION: Synthetic sequence designated GST-END
134 <400> SEQUENCE: 11
135 gaggcagatc gtcagtca          18
138 <210> SEQ ID NO: 12
139 <211> LENGTH: 16
140 <212> TYPE: DNA
141 <213> ORGANISM: Artificial
143 <220> FEATURE:
144 <223> OTHER INFORMATION: Synthetic sequence designated as BglI linker

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/821,160

DATE: 04/10/2001

TIME: 15:08:23

Input Set : A:\PTO.txt

Output Set: N:\CRF3\04102001\I821160.raw

```

146 <400> SEQUENCE: 12
147 aattcgccag gcaggc 16
150 <210> SEQ ID NO: 13
151 <211> LENGTH: 16
152 <212> TYPE: DNA
153 <213> ORGANISM: Artificial
155 <220> FEATURE:
156 <223> OTHER INFORMATION: Synthetic sequence designated at BglI linker
158 <400> SEQUENCE: 13
159 tcgagcctgc ctggcg 16
162 <210> SEQ ID NO: 14
163 <211> LENGTH: 101
164 <212> TYPE: DNA
165 <213> ORGANISM: Artificial
167 <220> FEATURE:
168 <223> OTHER INFORMATION: Synthetic sequence designated RSOL with DraIII site and
sequence
169 complementary to GST-EN
171 <220> FEATURE:
172 <221> NAME/KEY: N_region
173 <222> LOCATION: (24)..(84)
174 <223> OTHER INFORMATION: N refers to any nucleotide
177 <400> SEQUENCE: 14
178 atacacggcg tggctcttgca atannnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 60
180 nnnnnnnnnnn nnnnnnnnnnn nnttgactga cgatctgcct c 101
183 <210> SEQ ID NO: 15
184 <211> LENGTH: 23
185 <212> TYPE: DNA
186 <213> ORGANISM: Artificial
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Synthetic sequence including DraIII site designated RSF
191 <400> SEQUENCE: 15
192 atacacggcg tggctcttgca ata 23

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/821,160

DATE: 04/10/2001

TIME: 15:08:24

Input Set : A:\PTO.txt

Output Set: N:\CRF3\04102001\I821160.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14

L:180 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14